Article Title: Enhancing social tagging with automated keyword suggestions from the Dewey Decimal Classification

Author Details

Author 1 Name: Koraljka Golub
Department: UKOLN
University/Institution: University of Bath
Town/City: Bath
State (US only):
Country: United Kingdom

Author 2 Name: Marianne Lykke
Department: Department of Communication and Psychology
University/Institution: University of Aalborg
Town/City: Aalborg
State (US only):
Country: Denmark

Author 3 Name: Douglas Tudhope
Department: Faculty of Computing, Engineering and Science
University/Institution: University of South Wales
Town/City: Pontypridd
State (US only):
Country: United Kingdom

Acknowledgments:
This work in part results from a project called EnTag which was funded by UK's Joint Information Systems Committee (JISC) under its Capital Programme, for a period of 12 months in 2007 and 2008. Many thanks to Intute who supplied an extract of their social science collection, and to Diane Vizine-Goetz, Andrew Houghton, Michael Panzer of OCLC who provided the Dewey Decimal Classification and their assistance during the project.

Biographical Details:
Koraljka Golub is a research officer at UKOLN, University of Bath. Her major areas of interest are knowledge organization and information retrieval. More specifically, her projects deal with knowledge organization systems in contexts of social tagging, terminology registries, automated subject classification and indexing as well as evaluation methodology thereof.

Marianne Lykke is a professor at the Department of Communication and Psychology – University of Aalborg. Her areas of expertise include digital management, information seeking and information behaviour, and knowledge organization and information architecture. She is currently working with three major projects, one dealing with semantic components for facilitating access to domain-specific documents, another on evaluation of metadata and a third one on query modification and structuring.

Douglas Tudhope is a professor at the Faculty of Computing, Engineering and Science, University of South Wales. His main research interests are in knowledge organization systems and services. He is interested in how semantically indexed hypermedia systems can assist interactive and automatic retrieval. Current work investigates the application of semantic technologies and linked data to archaeology and cultural heritage generally.

Structured Abstract:

Purpose To explore the potential of applying the Dewey Decimal Classification as an established knowledge organisation system for enhancing social tagging, with the ultimate purpose of improving subject indexing and information retrieval.
Design/methodology/approach Over 11,000 Intute metadata records in politics were used. 28 politics students were each given 4 tasks, in which a total of 60 resources were tagged in two different configurations, one with uncontrolled social tags only and another with uncontrolled social tags as well as suggestions from a controlled vocabulary. The controlled vocabulary was Dewey Decimal Classification (DDC) comprising also mappings from the Library of Congress Subject Headings (LCSH).
Findings The results demonstrate the importance of controlled vocabulary suggestions for indexing and retrieval: to help produce ideas of which tags to use, to make it easier to find focus for the tagging, to ensure consistency and to increase the number of access points in retrieval. The value and usefulness of the suggestions proved to be dependent on the quality of the suggestions, both as to conceptual relevance to the user and as to appropriateness of the terminology.
Originality/value No research has investigated the enhancement of social tagging with suggestions from the Dewey Decimal Classification, an established knowledge organisation system, in a user trial, comparing social tagging only and social tagging enhanced with the suggestions. This paper is a final reflection on all aspects of the study.

Keywords:
Social tagging, subject indexing, information retrieval, controlled vocabulary, knowledge organisation system, Dewey Decimal Classification (DDC), Library of Congress Subject Headings (LCSH)

Article Classification:
Research paper

Enhancing social tagging with automated keyword suggestions from the Dewey Decimal Classification

1 Introduction

Knowledge organization systems (in further text: KOS) such as classification schemes and subject heading systems play a crucial role in information discovery. They help reduce the ambiguity of end users' natural language during search and improve precision and recall of the retrieved resource set. However, although they have been used as tools for information organization and retrieval for over a century now, there are several issues of applying them today; the major ones are the associated cost, slowness to accommodate new concepts, and traditionally a resource-oriented rather than user-oriented approach to indexing.

Social tagging by end users, resulting in collections of terms popularly called folksonomies, has the potential to help alleviate these issues. It is much cheaper to produce as it is done by end-users who need not be paid or educated on how to index. Creating KOS themselves is not required; new terms and concepts are easily added as end users find them appropriate; and, it is user-oriented in that it is the end users who assign the tags that are of use to them, and they include tags and concepts that may not be found in KOS at hand.

However, this potential is seriously questioned because the majority of existing social tagging applications and the resulting folksonomies suffer from a lack of even basic word form control, whereas mechanisms such as those for regulating homonymy, polysemy and synonymy typical of KOS are entirely absent.

Because both traditional KOS and folksonomies have their advantages and disadvantages, the study presented here investigated the potential for enhancing social tagging with a KOS, with a view to improving the quality of tags for increased information discovery and retrieval. The objectives were the following:

1. Investigate indexing aspects when using only social tagging versus when using social tagging in combination with a KOS;
2. Investigate the influence on retrieval of only social tagging versus social tagging in combination with a KOS.

Early reports were reported at conferences as the project progressed (Golub *et al.,* 2008a; Golub *et al.,* 2009a; Golub *et al.,* 2009b; Lykke *et al.,* 2011; Matthews *et al.,* 2010). The project explored and compared two different contexts: 1) tagging by readers using the Intute digital collection and the Dewey Decimal Classification (DDC), and 2) tagging by authors submitting papers to the Science and Technology Facilities Council (STFC) repository and the ACM Computing Classification Scheme. This final paper is an in-depth analysis of the Intute study and the DDC context and provides new findings with a comprehensive overview.

The remainder of this paper provides Background information in section 2 which includes rationale and related work, followed by a section on the Methodology. Section 4 presents the Results which are followed by Conclusions in section 5.

2 Background

2.1 Rationale

KOS such as classification schemes, thesauri and subject heading systems have mechanisms that help translate the natural language of authors, indexers, and users into a vocabulary used for indexing and retrieval. These mechanisms ensure consistency through uniformity in term format and the assignment of terms, indicate semantic relationships among terms, support browsing by providing consistent and clear hierarchies in a navigation system and support retrieval (National Information Standards

Organization, 2005). KOS have been used in libraries and databases to support information retrieval, some since the 19th century.

Major issues with KOS which are of relevance here follow:

1) There is a major cost involved in the production of a KOS or in acquiring access to an existing one, to start with;
2) The associated cost of controlled subject indexing is high, especially in the light of the ever-increasing numbers of available digital resources that need to be catalogued and classified or indexed;
3) Editing the existing KOS for inclusion of new terms and concepts can be too slow of a process, preventing end users from retrieving resources using these new concepts and terms;
4) While they have been built with the end-user in mind, many KOS are largely based on literary warrant, making them less efficient in end-user search and retrieval; and,
5) In spite of the fact that user-oriented subject indexing and classification is encouraged, the more traditional resource-oriented approach is often followed in real-life resource collections.

Folksonomies are popularly understood to be the products of social tagging, the latter referring to end-users assigning keywords to their chosen resources in Web-based services such as del.icio.us (2011) (resources being bookmarks) or Flickr (2011) (resources being photographs). Although folksonomies and social tagging are often attributed to have appeared through collaboration, Spiteri (2007) points out that, while a person may access and use tags assigned by others he or she may not actively collaborate in the creation and assignation of tags. Web-based services that enable their users to tag began to proliferate in the mid-2000s; today tagging is also part of other Web 2.0 applications, such as blogs and micro-blogs.

Social tagging has the potential to help alleviate some of the issues pertinent to KOS as follows:

1) It is much cheaper to produce as it is done by end-users who are need not be paid or educated on how to index, and creating the KOS themselves is not required;
2) New terms and concepts are easily added as end users find them appropriate;
3) It is user-oriented in that it is the end users who assign the tags that are of use to them, and they include tags and concepts that may not be found in KOS at hand.

However, in contrast to established (inter)national standards and indexing policies followed by information experts assigning controlled terms, it is more or less non-transparent why and how users tag. Furner (2010, p. 1862) identifies potential goals for tagging to be 1) knowledge-oriented goals (e.g., to improve resource discovery), 2) self-oriented goals (e.g., to attract attention to the resources they have created or collected), 3) socially-oriented goals (e.g., to help people share resources and knowledge about resources and collaborate with others), and 4) institutionally-oriented goals (e.g., to help the institution become visible in its community). Furthermore, unlike trained information professionals who possess high levels of indexing expertise gained through specialised university-level education, social taggers have little indexing expertise. In social tagging applications there is no requirement to choose the ‘right’ tag or even the ‘right’ form of the term or to wait until the tag is approved by the KOS editors.

As a result, tags in existing tagging applications suffer from a lack of even basic word form control like spelling mistakes, different ways of writing the same tag (‘redwine’, ‘red_wine’, ‘red wine’), different word forms such as singular or plural and spelling variations. Mechanisms such as those for regulating homonymy, polysemy, synonymy and for expressing other semantic relationships typical of KOS are entirely absent. Consistency in specificity and exhaustivity of indexing which are defined by indexing policies where KOS are used is also non-existent (e.g., the same resource is tagged by both ‘river’ and ‘Danube’; 20 tags are added to one resource but only 1 tag to another). In addition, only a proportion of resources are tagged exhaustively; research has shown that resources which have received most tags in the past are most likely to receive tags in the future (Furner, 2010). Also, although the majority of tags are subject-related, genre, format, author and other tags are assigned (e.g., Catarino and Baptista, 2008), including tags for exclusive personal use (‘toread’, ‘myparty’, ‘me’ etc.).

Thus, “the cost savings made in the provision of low-quality indexing are cancelled out by the high costs incurred by searchers who fail either to find everything that they want (low recall) or, often more frustratingly, to avoid everything that they do not want (low precision)” (Furner, 2010, p. 1861). This being the case, end users themselves may appreciate tagging guidelines and controlled form suggestions even if they are more time consuming (cf. results of our study below). Proposals of methods for improving the quality of tagging so far include providing taggers with ways of recording explanations or justifications of their tagging decisions, allowing taggers to give feedback on the decisions made by others, automatically suggesting standard forms of tags, and automatically suggesting different facets of resources for taggers to describe (Furner 2010, 1863), the latter two of which are explored in our study based on prominent a KOS.

2.2 Related work

The need for KOS in relation to folksonomies has been reported in the literature. Weller (2007) compared ontologies and folksonomies, suggesting that they are not to be seen as rivals but complementary to each other. Syn and Spring (2013) show that social tags can be used as metadata and present metrics to select the tags which have that potential. Noruzi (2007) gave seven arguments for why a folksonomy-based system should use a thesaurus, emphasizing that it is impossible to maintain consistency over time or across folksonomy users without a thesaurus. Spiteri (2007) analyzed tags from three social tagging services as to their conformance to a US set of indexing guidelines. The study showed how ambiguity and polysemy are indeed problems in the structure of the folksonomy tags, although homographs and ambiguous tags each constituted less than one-quarter of the tags in each of the three services. She encouraged library catalogues be made to support social tagging, but advised that guidelines for the choice and form of tags are provided.

A number of attempts at tag control exist. ZigTag (http://zigtag.wordpress.com) used to be a social bookmarking service that provides disambiguating definitions of entered terms from which the user can choose when tagging or searching. Marchetti et al. (2007) describe the SemKey system in which they use WordNet and Wikipedia to introduce unique concepts. They conclude that there are problems with both: WordNet contains parts of speech information and structured network of relations between them, but lacks data for proper names disambiguation; Wikipedia on the other hand has strong proper names coverage but lacks a structured set of relations between the concepts described. FaceTag (http://www.facetag.org/), a prototype of a collaborative tagging tool for bookmarking information architecture resources, attempts to combine user-generated tags with a richer faceted classification scheme. Faviki (http://faviki.com/) is another social bookmarking tool that uses DBpedia, Wikipedia-based RDF data, as unique concepts. For Connotea (http://www.connotea.org/, discontinued on 12 March 2013), a service for organizing references, an add-on tool called Entity describer started being developed at one point, which was to allow taggers to select terms from a pre-existing list, which could well have been a KOS. Also, much of the content of Connotea consists of references imported from bibliographic databases which carry metadata assigned from a KOS used in those databases. Of these tools and services, none, however, offer or actually offered a KOS as a pool from which to choose tags.

Existing services that allow both KOS terms and social tags to co-exist include the following. Of library catalogues, PennTags at the University of Pennsylvania (2005), Ann Arbour District Library (2011), Opacial (‘OPAC social’) of the Greek Panteion University (2008) all allow end users to add tags to existing catalogue records already containing controlled subject headings. Montana State University’s Electronic Thesis and Dissertations collection (2011) also allows subject headings as well as end-user tags and supports browsing by tags and by subjects. Library vendors such as Ex Libris’ Primo (2011) and Innovative Interfaces’ Encore (2011) have included tagging and social participation. Library of Congress makes available some of its image collections on Flickr and allows end users to contribute with tags (Raymond 2008). Similarly, the Picture Australia project by the National Library of Australia (2010) encourages end users to upload their pictures related to Australia on Flickr and add tags there. In mobile phones, ZoneTag is an application for images which can be uploaded to Flickr, at which point tags suggestions are provided (Naaman and Nair 2008). Still, none of the existing services allow end-users to choose from KOS terms as tags. Hayman (2007) developed a proof of concept whereby users could tag resources by choosing from an established KOS or by entering their own terms; users’ own terms are used later to feed back into the taxonomy to improve its quality. Rolla (2009) compared the LibraryThing (a social service for books cataloguing) tags for a group of books and the library-supplied subject headings (Library of Congress Subject Headings – LCSH) for the same books, concluding that, since users and cataloguers describe material from different aspects, user

tags can enhance subject access to library materials, but cannot replace KOS. Yi and Chan (2009) matched a sample of tags from Delicious, a social bookmarking service, also against LCSH, showing that two thirds of the tags matched LCSH, implying that the other third may provide additional access points.

As known to the authors, no research has investigated the enhancement of social tagging with suggestions from a KOS in a user evaluation with an existing retrieval application, comparing simple social tagging and tagging enhanced with KOS, and looked into implications for indexing and retrieval thereof, and users' opinions about it.

3 Methodology

3.1 Data collection, KOS and demonstrator

The Intute subject gateway is a quality-controlled UK-based collection of hand selected and catalogued Web resources for education and research (it was maintained and updated until July 2011 and is still available as of 15 August 2012). For each resource the metadata (or catalogue) record includes title, description, controlled keywords, uncontrolled keywords, classification, type, URL, format and other elements. In our study, 11,042 Intute metadata records were chosen, all from the area of politics. The selected records were indexed ('controlled keywords' field) by IBSS thesaurus (IBSS: International Bibliography of the Social Sciences, by ProQuest (2011)) and by HASSET thesaurus (Humanities and Social Science Electronic Thesaurus, by UK Data Archive (2011)). Uncontrolled keywords mostly comprise geographical names.

The KOS from which tagging suggestions are provided in the demonstrator is the Dewey Decimal Classification (DDC) (OCLC, 2011), including DDC captions (names of alpha-numerical classes), DDC Relative Index terms pointing to the captions, and mappings from Library of Congress Subject Headings (LCSH). Thus, the suggestions offered come from the three sources. The choice of DDC was due to several factors: it offers subject-based hierarchical browsing of its captions and classes; the captions often comprise complex terms which denote a concept very specifically, i.e., to a great detail; it is a very widespread classification scheme in use and many potential users of a tagging service may already be accustomed to it.

The demonstrator (Moon *et al.*, 2008) is a .NET application, employing Visual Studio, C# and ASP .NET. It operates over an extract of the Intute metadata records held on the local server. The URL is provided so that the participant can view the resource, along with the title and the description fields. Original controlled keywords, uncontrolled keywords and classification fields have been stripped out for purposes of the study.

The demonstrator comprises three major interfaces: searching, simple tagging, and enhanced tagging. Once a person logs in, she arrives to the searching interface (Figure 1). The searching interface provides the following features:

1. Main tag cloud: a tag cloud with tags linked to resources to which they were assigned. It is an alphabetical list of all tags in the demonstrator, with different font sizes relative to popularity. Filter By drop-down menu on top offers the My Tags option which presents the current taggers's tags only. By default everyone's tags are shown (Everyone's Tags).
2. Taggers: a cloud of names of taggers linked to resources they indexed.
3. A free-text search box, with an option to limit searching to tags, title and description fields (Search For Documents).

The resources found are shown in the Results pane. They are automatically ranked according to the MySQL full-text natural language search. Title, description and existing tags are shown here. By clicking on the Title, the URL opens in a new window.

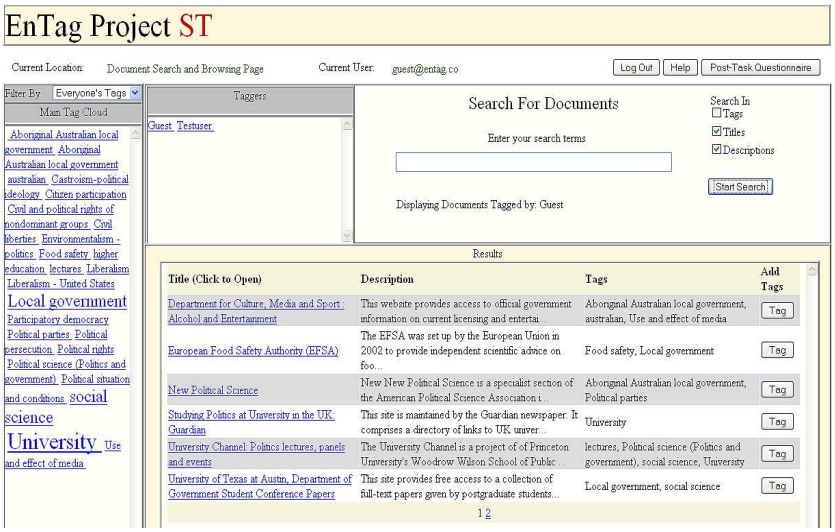

Figure 1. Searching interface

Once a resource is selected from search results (the Results pane), clicking on the "Tag" button will return a tagging interface (Figure 2). Here title, URL, and description are displayed. Two tagging interfaces are provided: 1) the Simple Tagger, with tagging features common in popular social tagging applications; and 2) the Enhanced Tagger, with additional suggestions from the KOS. In the study, the log-on screen provides the choice of the interfaces. Both tagging interfaces have the following options from which to select tags:

1. Main tag cloud.
2. Taggers: names of taggers linked to tags they have used, the latter listed in the All {Tagger Name}'s Tags pane; and,
3. My Tags For This Document.

By clicking on a selected tag in any of these options, the tag will be shown in the text box (under the Document Details pane). By pressing the Tag Document button the tag will be added to the resource as well as listed in the "My Tags for This Document" pane. A tag can also be typed in.

The Enhanced Tagger (Figure 2) additionally provides suggestions from the KOS that are presented in three panes at the bottom of the screen. In the first pane to the left ("Automatically suggested matches, Find appropriate context(s)"), DDC captions are listed. They are automatically derived by a string-matching comparison of DDC captions to a user-entered term from the text box upon clicking the Suggest button. Immediately after the user comes to the enhanced tagging page, initial suggestions are listed. These are automatically generated by treating the resource's title as if it had been entered as a tag.

If the user clicks on one of the listed captions from the first pane, its narrower and broader classes are shown in the second pane ("Explore hierarchy around the selected context"), allowing interactive browsing of the hierarchical context. Simultaneously, in the third pane ("Select/edit relevant tags") a tag-cloud-like list of DDC captions, DDC relative index terms and LCSH mapped terms is presented as a source of suggestions from which the user may select a tag. Selecting a tag copies it to the text box, where it can be further edited; pressing the Tag Document button adds the tag to the resource.

Figure 2. Tagging interface: Enhanced Tagger

What follows is an example of how a tag is chosen from suggestions in the Enhanced Tagger (see Figure 2). After searching for the term 'slavery' in all fields, a list of resources is returned. The resource chosen is 'Slavery in New York', an online exhibition on history of slavery in New York State (http://www.slaveryinnewyork.org/). By clicking on the "Tag" button, the enhanced tagging interface opens. In the text box in the middle of the screen the tag 'slavery' is entered and the Suggest button clicked, to get DDC-based suggestions. The first bottom pane displays 19 matching classes. The top two of these belong to political science. The End of Political Science line means that the classes below it belong to different subject areas which could be related to politics. Mousing over each caption will display its immediate superordinate class. The top found classes are the following (*the superordinate class in parentheses*):

Slavery and emancipation *(Political science)*
Emancipation *(Slavery)*
----- *End of political science* -----
Discriminatory practices and slavery *(Ethics of social relations)*
Extension of slavery *(Administration of Abraham Lincoln)*
Slavery *(Systems of labor in Economic institutions)*
American fiction-1830-1861, ... *(American literature)*
Etc.

3.2 The user study set-up

3.2.1 Participants

There were 66 participants who registered for the study. Six of them took part in pilot testing. Of the remaining 60, 54 went on to complete the pre-study questionnaire, and 42 who also did some tagging. Twenty-eight participants completed all study tasks and the analysis and results reported here are based on their input.

The 28 participants were all political science students at British universities, with one from the European University Institute in Italy. The participants were recruited via a written call for participation which was emailed to mailing lists for political science students as well as to several dozens top-rated political science departments in the UK sent out throughout the UK. In addition, paper adverts were put up at the University of Bath and a live presentation was held for a class of political science students at Exeter University; these two efforts, however, yielded only one student.

Each participant who completed the study received a £50 Amazon voucher. The study took place remotely at participants' homes using their own computers, supported by communication via email.

Based on the pre-study questionnaire, it was shown that the majority of the participants have had solid political science experience and were experienced Web users. Half of them have used tagging applications before but conducted little tagging. Detailed characteristics are as follows:

1. The gender distribution was equal.
2. The majority were aged between 21 and 25 (16 of them; 5 were aged 26-30; 4 were 20 or younger; 2 were 31-35, and 1 was 36 or older).
3. The majority had English as their native language (20 of them).
4. The majority had solid political science knowledge, as indicated by the years spent studying political science: 18 have been studying political science for over 2 years (over 5 years: 5 of them; 4-5 years: 3; 3-4 years: 4; 2-3 years: 6; 1-2 years: 8; less than 1 year: 2).
5. They came from 13 British universities and 1 came from the European University Institute: University of Essex (5 of them), University of Birmingham (4), University of Leeds (4), Birkbeck College University of London (2), London School of Economics (2), Queen Mary University of London (2), University of East Anglia (2), University of Manchester (2), Kings College Cambridge (1), University of Bristol (1), University of Exeter (1), University of Newcastle (1), University of Warwick (1); European University Institute (1).
6. The majority were experienced Web users, as judged by the number of years spent using the Web – 27 have been using the Web for over 4 years (7 or more years: 21 of them; between 4 and 6 years: 6; between 1 and 3 years: 1);
7. The majority (23) have never used Intute before (never: 23 of them; once or twice a year: 4; once or twice a month: 1).
8. Approximately half have used tagging applications before (16 of them); the tagging applications they used were: Flickr: 8 of them, Last.fm: 4, Del.icio.us: 1, Technocrati: 1, Other: 2.
9. The ones who used tagging applications have relatively little tagging experience, as judged from the number of resources they tagged in those applications (20 or less resources tagged: 11 of them; 100-200 resources tagged: 2; more than 200 resources tagged: 1).
10. Almost a third had some acquaintance with DDC (8 of them), and none had any acquaintance with LCSH or other KOS.

3.2.2 Resources and tasks

Each participant was given 4 tasks, and in each task 15 resources were to be tagged. Each task covered one topic of relevance to political science students. Topic-wise, two tasks were controlled, one in each interface, and two tasks were free, also one in each interface. Two were to be completed in the Simple Tagger, and two in the Enhanced Tagger. In order to reduce the learning influence on the results, the tasks were rotated as follows:

1. First rotation: 1. controlled task, Simple Tagger, 2. controlled task, Enhanced Tagger, 3. free task, Simple Tagger, 4. free task, Enhanced Tagger;
2. Second rotation: 1. controlled task, Enhanced Tagger, 2. controlled task, Simple Tagger, 3. free task, Simple Tagger, 4. free task, Enhanced Tagger;
3. Third rotation: 1. controlled task, Simple Tagger, 2. controlled task, Enhanced Tagger, 3. free task, Enhanced Tagger, 4. free task, Simple Tagger; and,
4. Fourth rotation: 1. controlled task, Enhanced Tagger, 2. controlled task, Simple Tagger, 3. free task, Enhanced Tagger, 4. free task, Simple Tagger.

The first participant was given the first rotation of tasks, the second participant the second rotation etc.

In each task the first instruction was to search for documents and then tag 15 of them. In controlled tasks exact search phrases were given and the top 15 retrieved documents were to be tagged, so that the same set of documents was retrieved and tagged each time. In free tasks they could choose any 15 documents to tag that they found relevant.

In case a URL had been unavailable, the instruction was to move on to the following one. The analysis showed that in controlled tasks 53 documents were tagged at least once in the Enhanced Tagger, and 41 in the Simple Tagger (instead of 15 in each). This is because participants did not closely follow the ‘top 15’ instruction and because the URLs of some of the top 15 retrieved documents were temporarily unavailable; in two cases in the Simple Tagger the participants also chose a different search phrase (‘EU

integration' and 'european AND integration' rather than 'european integration' that was specified) which ordered the retrieved set of documents differently.

Tagging guidelines were to spend between 5 and 10 minutes on each document and to tag as many aspects and topics the participant considered appropriate for the task. The participant was instructed to open the URL, but did not need to follow further links. The log analysis showed that the URL was indeed followed for nearly every resource by every participant, and some URLs were followed more than once by the same participant. In case of long documents the guideline was to focus on the abstract, introduction, conclusion, headings and table of contents. When tasks were being completed in the enhanced interface, the instruction was to try to consider the tagging suggestions if appropriate with no requirement to actually use any of them.

Topics for the controlled tasks were suggested by a subject expert, a PhD student in political science, who also confirmed that there were at least 20 resources in the database relevant to the topics. The controlled task for the Simple Tagger was on the topic of European integration:

*Imagine that as part of one of your courses, you are asked to write a four-page essay on the topic of **European integration**, as a joint project in groups of four. The essay should critically discuss existing theories about the creation of the European Union and its institutions. Your lecturer has instructed you to look for resources in the EnTag system. Since you will be working together with three other students, you should tag the documents you retrieve with tags that would be useful to you but would also enable other students to find those documents in EnTag and understand from your tags what the documents are about.*

The controlled task for the Enhanced Tagger was on the topic of peacekeeping:

*Imagine that as part of one of your courses, you are asked to write a four-page essay on the topic of **peacekeeping**, as a joint project in groups of four. The essay should describe and discuss deployment of non-partisan military forces to separate two sides of a conflict that have already agreed on an armistice (are no more at war with each other). Your lecturer has instructed you to look for resources in the EnTag system. Since you will be working together with three other students, you should tag the documents you retrieve with tags that would be useful to you but would also enable other students to find those documents in EnTag and understand from your tags what the documents are about.*

3.2.3 The flow of the study

After signing the participation form and completing the pre-study questionnaire, the participants received the instructions which included technical settings such as enabling scripting and zooming screen display. Before starting with the tasks, each participant was to get acquainted with the system which is why several brief exercises were given. After each task the participant was to complete a post-task questionnaire, and after finishing all the tasks a post-study questionnaire. All of these documents are available (Golub *et al.,* 2008b).

3.2.4 Collecting the data

The main method of data collection was logging the steps the participants took in the demonstrator. In order to help contextualize the results, pre-study, post-task, and post-study questionnaires were used as described above. This helped us collect data from different angles and derive more confident conclusions. The data were collected in the period between 25 June and 21 July 2008.

Table 1 shows an excerpt from the log data for one particular user, for one resource, in the Enhanced Tagger. She logged on, searched for 'peacekeeping' and opened one of the retrieved resources. The resource is a report about civil-military relations in Afghanistan and Liberia from 2006 titled 'Principles and pragmatism: Civil military action in Afghanistan and Liberia' found at Cordaid, a Dutch Christian agency (the resource available at http://www.cordaid.nl/nl/cmrreport.pdf).

4-5 She entered the term 'NGO' (standing for non-governmental organization) and clicked on the Suggest button to get DDC suggestions. She decided to add the tag 'NGO' to the resource (in DDC there are no matches to NGO).

6-11 She entered the term 'civil-military relations' and clicked on the Suggest button again. From the list of suggested classes (left bottom pane) she chose 'Foreign policy and specific topics in international relations'. Then she browsed around its hierarchy and chose its narrower term 'International conflict'

(middle bottom pane). From the cloud of controlled terms (right bottom pane), she chose 'Conflict – international relations' and added it as another tag for the resource. There she also clicked on 'Foreign policy and specific topics in international relations' but chose not to assign it.

12-18 She repeated the cycle with another term: she entered 'liberia' and asked for DDC suggestions. In this case the first pane returned 'Liberia' as the top class and the hierarchy shown for it in the middle pane is for 'Liberia' where she chose the after 1945 context. She then clicked on three suggested tags from the third pane and added two of them: 'Liberia – 20th century' and 'Liberia - History - Civil War, 1989-'.

19-24 She repeated the cycle with yet another term: she entered 'afghanistan' and asked for DDC suggestions. In this case the first pane also returned 'Afghanistan' as the top class and the hierarchy shown for it in the middle pane is for 'Afghanistan' where she first chose the after 1919 context, and then within it the after 2001 context. She clicked on 'Afghan War, 2001' from the third pane and added it to the resource. She clicked again on 'Afghan War, 2001' from the third pane.

25-31 She then entered 'isaf' (for which there are no matches in DDC) and added 'isaf', 'UNML', and 'NATO' directly. Of these three acronyms 'NATO' does return matches in DDC and she then went on to explore its context and added 'North Atlantic Treaty Organization' from the third pane of suggested terms.

32-35 She asked for suggestions for 'development' and then added it as the tag, without first exploring its context in DDC for a more precise match. She also added 'cordaid' and found no matches for it in DDC.

36-42 Finally, she searched for 'civil society', chose 'Armed services' from the list of retrieved DDC matches and added as tags 'Armed Forces – Political activity' and 'Civil supremacy over the military'. She then returned to the searching page.

Table 1. An excerpt from log data

Step	Type of activity	Term being considered	Tag assigned
1	Log on		
2	Document search ('peacekeping')		
3	Document opened		
4	DDC Suggest button clicked	NGO	
5	Tag added		NGO
6	DDC Suggest button clicked	civil-military relations	
7	DDC treeview clicked (first pane)	Foreign policy and specific topics in international relations	
8	DDC hierarchy clicked (middle pane)	International conflict	
9	DDC suggestion cloud clicked (third pane)	Conflict - international politics	
10	Tag added		Conflict - international politics
11	DDC suggestion cloud clicked (third pane)	Foreign policy and specific topics in international relations	
12	DDC Suggest button clicked	Liberia	
13	DDC hierarchy clicked (middle pane)	1945-	
14	DDC suggestion cloud clicked (third pane)	Liberia - History - Civil War, 1989- - Peace	
15	DDC suggestion cloud clicked (third pane)	Liberia - 20th century	
16	Tag added		Liberia - 20th century
17	DDC suggestion cloud clicked (third pane)	Liberia - History - Civil War, 1989-	
18	Tag added		Liberia - History - Civil War, 1989-
19	DDC Suggest button clicked	Afghanistan	
20	DDC hierarchy clicked (middle pane)	1919-	
21	DDC hierarchy clicked (middle pane)	2001-	
22	DDC suggestion cloud clicked (third pane)	Afghan War, 2001-	
23	Tag added		Afghan War, 2001-

24	DDC suggestion cloud clicked (third pane)	Afghan War, 2001-	
25	DDC Suggest button clicked	Isaf	
26	Tag added		isaf
27	Tag added		UNMIL
28	Tag added		NATO
29	DDC Suggest button clicked	nato	
30	DDC suggestion cloud clicked (third pane)	North Atlantic Treaty Organization	
31	Tag added		North Atlantic Treaty Organization
32	DDC Suggest button clicked	development	
33	Tag added		development
34	Tag added		cordaid
35	DDC Suggest button clicked	cordaid	
36	DDC Suggest button clicked	civil society	
37	DDC suggest treeview clicked (first pane)	Armed services	
38	DDC suggestion cloud clicked (third pane)	Armed Forces - Political activity	
39	Tag added		Armed Forces - Political activity
40	DDC suggestion cloud clicked (third pane)	Civil supremacy over the military	
41	Tag added		Civil supremacy over the military
42	Go to searching		

3.2.5 Analyzing the data

Based on the log analysis the following results are derived: number of tags assigned; distribution of tags; where the participants browsed for tags; where the tags are taken from; and, retrieval implications based on overlap of simple and enhanced tags with existing resources' metadata records, overlap of search terms with assigned tags versus overlap of search terms with existing metadata.

The post-task and post-study questionnaires helped provide the context for the results derived from the log analysis. More specifically, the post-task questionnaire provided insight into the participants' familiarity with task topics, easiness, satisfaction and certainty in the choice of tags for the task, and, usefulness of the offered tagging-related features for the task. The post-study questionnaire helped indicate the degree to which the participants enjoyed the study, the ease of learning and using the two interfaces, usefulness of the system in real life, what is liked and what is disliked about the two interfaces as well as what general comments and suggestions were.

In addition, overlap of simple and enhanced terms with title terms and with existing keywords from the metadata records was qualitatively analyzed. From the pool of 21 resources which were tagged in both interfaces 12 resources were randomly selected. The tags were first compared against the reference standard which was taken to be the existing controlled and uncontrolled keywords in the metadata records whereby seven different types of matches were identified. These were defined as follows:

1. Title term, where the tag is identical to title disregarding the order of terms;
2. Preferred term, where the tag is identical to one of the (un)controlled keywords;
3. Synonym term, where the tag has the same meaning as one of the (un)controlled keywords;
4. Narrower term, where the tag is hierarchically subordinate to one of the (un)controlled keywords;
5. Broader terms, where the tag is hierarchically superordinate to one of the (un)controlled keywords;
6. Related term, where the tag is associatively related to one of the (un)controlled keywords; and,
7. Spelling mistakes, where the tag is a spelling mistake.

Each tag was coded as one match type only, in the order given above, i.e., if more than one matches were found then the one higher up in the above list was assigned. The coding was carried out by three researchers separately and then harmonized. See (Lykke *et al.*, 2011) for further details.

4 Results

4.1 Number of tags

As seen from Table 2, in total 7568 tags were assigned in both interfaces and in all the tasks. More tags were assigned in the Simple Tagger than in the Enhanced Tagger, which could be explained by the fact that exploring KOS suggestions in the Enhanced Tagger requires more effort and hence as a result fewer tags are assigned. Each participant assigned in total an average of 278 tags to 60 documents. A few more tags were assigned in the Simple Tagger and a few more in the free task, probably due to faster tagging process in the Simple Tagger and more motivation in the free task. For the 94 resources that were tagged in controlled tasks in total, on average 41 tags per resource were assigned, and for the 751 resources that were tagged in free tasks, on average 5 tags per resource were assigned. The latter figure is similar to findings of Angus *et al.* (2008) where on average four tags were assigned in Flickr.

Table 2. Number of tags

	Simple Tagger	**Enhanced Tagger**	**Total**
Tags in total	4022	3546	7568
Controlled task	2025	1688	3713
Free task	1997	1858	3855
Tags per resource (controlled)	average 49 (41 resources)	average 32 (53 resources)	average 41 (94 resources)
Tags per resource (free)	average 5 (374 resources)	average 5 (377 resources)	average 5 (751 resources)
Tags per tagger (controlled)	average 72	average 63	total average 135
Tags per tagger (free)	average 74	average 69	total average 143

Table 3 shows the distribution of tags, which follows the power law in both interfaces, i.e., a few tags are used by many and many tags are used by a few, which is common in tagging (Furner, 2010). There are over 2000 tags in each interface which are assigned between 1 and 5 times and there are 17 tags in each interface which are assigned 16 or more times.

Table 3. Distribution of tags by usage

Number of times a tag is assigned	**Number of tags assigned in Simple Tagger**	**Number of tags assigned in Enhanced Tagger**
1-5	2250	2036
6-10	92	70
11-15	24	20
16 and above	17	17

4.2 Choosing a tag

As seen from Table 4, in the Simple Tagger tasks the majority of tags (90.9%) are assigned by typing them directly in (Typing own tag), as common in social tagging applications. Of other sources, another tagger's tags are chosen in 6.8% of cases, and tags from the main tag cloud in 2.3% cases. In the Enhanced Tagger 71.2% tags are typed in, while 16.9% come from DDC-based suggestions (DDC tag). From other sources, another tagger's tag (8.5% of tags) and the main tag cloud (2.5% of tags) were used. The latter figure is rather small considering the frequency of the feature in tagging applications. Also, 0.9% tags were selected from own tags.

Table 4. Choosing a tag

Activity	**Simple Tagger**		**Enhanced Tagger**		**Total**	
Typing own tag	3656	90.9%	2525	71.2%	6181	81.7%
Main tag cloud	94	2.3%	88	2.5%	182	2.4%
Own tag	0	0%	32	0.9%	32	0.4%

Certain tagger's tag	272	6.8%	303	8.5%	575	7.6%
DDC tag			598	16.9%	598	7.9%
Total	4022	100%	3546	100%	7568	100%

Table 5 shows where the participants looked for potential tags (not necessarily selecting the tag). In the Simple Tagger the most frequently visited feature was others' tags, 72.9% in total (Certain Tagger's Name (18.9%) and Certain tagger's tag (54.0%)). Main tag cloud and Own tag were less used (16.8% and 10.3% respectively). In the Enhanced Tagger most frequent activities were the ones related to exploring DDC suggestions, 74.0% in total, with the DDC Hierarchy used the least. Of other features in the Enhanced Tagger others' tags were most popular, 18.2% in total, as in the Simple Tagger. In comparison, post-task questionnaires show that the participants found the following features useful: list of own tags, DDC treeview (first pane) and DDC suggested tags (third pane). Main tag cloud was considered neither useful nor not useful; and the same was thought of others' names and DDC middle pane. Others' names were considered not useful in free tasks which is understandable because there were only 28 taggers and each chose a different free task.

Table 5. Browsing for tags

Activity	**Simple Tagger N = 614**	**Enhanced Tagger N = 2468**
Main tag cloud clicked	16.8%	5.1%
Own tag clicked	10.3%	2.7%
Certain tagger's name clicked	18.9%	3.7%
Certain tagger's tag clicked	54.0%	14.5%
DDC Suggest button clicked		28.9%
DDC treeview clicked (first pane)		13.7%
DDC hierarchy clicked (second pane)		3.2%
DDC suggested tags clicked (third pane)		28.2%
	100.0%	100.0%

Below are two illustrative examples of tagging two resources by two of the participants, demonstrating the order of places consulted before choosing a tag. Each line starts with a step the participant took followed by colon and the tag that was considered. The vertical bar sign separates two steps when the steps relate to the same tag. The line that starts with 'Tag added' means the tag was typed in without any sources consulted. The line that starts with another type of activity lists all the steps taken before the tag was chosen; it ends with 'Tag added' and shows which tag was chosen.

In first example for the Simple Tagger (Example 1), the participant mostly added tags by typing them directly in (20 times); in two cases she modified the tags she had already entered by first removing them (steps 2 and 15); in one case she chose the tag from the main tag cloud (step 23). Notice the spelling mistakes in steps 11 and 17 ('intergration', 'Europes'), common in tagging applications of today, which may in real systems be largely prevented by implementing automated spelling correction and by providing suggestions from a KOS.

Example 1: the Simple Tagger

- for resource titled 'Centre for European Politics, Security and Integration' (http://www.ssees.ucl.ac.uk/cepsi.htm)

1. Tag added: Centre for European Politics, Security and Integration
2. Own tag clicked: Centre for European Politics, Security and Integration | Remove Tag: Centre for European Politics, Security and Integration | Tag added: Centre for European Politics, Security and Integration
3. Tag added: UCL
4. Tag added: School of Slavonic and East European Studies
5. Tag added: politics of contemporary Europe
6. Tag added: contemporary Europe
7. Tag added: transition, integration, and globalisation

8. Tag added: EU transition
9. Tag added: european transition
10. Tag added: European integration
11. Tag added: EU intergration
12. Tag added: globalisation
13. Tag added: sovereignty
14. Tag added: europe of the future; future Europe
15. Own tag clicked: europe of the future; future Europe | Tag removed: europe of the future; future Europe | Tag added: future Europe
16. Tag added: lectures
17. Tag added: Europes interaction with other polities
18. Tag added: International Relations
19. Tag added: Security Studies
20. Tag added: Comparative European Politics
21. Tag added: public policy
22. Tag added: Political Sociology
23. Main tag cloud clicked: University College London | Tag added: University College London

Example 2: the Enhanced Tagger

- for resource titled 'United Nations Peacekeeping Operations' (http://www.un.org/en/peacekeeping/)

1. DDC Suggest button clicked: United Nations | *Tag added*: United Nations
2. DDC Suggest button clicked: Peacekeeping Operations | *Tag added*: Peacekeeping Operations
3. DDC Suggest button clicked: armistice | *Tag added*: armistice
4. DDC Suggest button clicked: conflict | *Tag added*: conflict
5. DDC Suggest button clicked: international peace | DDC suggestion cloud clicked: International cooperation | *Tag added*: International cooperation
6. DDC suggestion cloud clicked: Peace-international politics | *Tag added*: Peace-international politics
7. DDC suggestion cloud clicked: Promotion of peace and international order | *Tag added*: Promotion of peace and international order
8. DDC Suggest button clicked: security | DDC Suggest TreeView clicked: International cooperation | *Tag added*: security
9. DDC suggestion cloud clicked: International arbitration-international politics | *Tag added*: International arbitration-international politics

In the second example for the Enhanced Tagger (Example 2), in 5 out of 9 cases, the participant entered her own tags and asked for DDC suggestions by clicking on the DDC Suggest button (steps 1, 2, 3, 4, 5). In three cases the participant chose the tag from the list of related tags after entering term 'international peace' ('International cooperation', 'Peace-international politics', 'Promotion of peace and international order' (steps 5, 6, 7)). Finally, the participant browsed the DDC treeview and decided to assign the tag she used initially ('security') as well as to assign one from the DDC suggestion cloud ('International arbitration-international politics') (steps 8 and 9).

The instructions for the Enhanced Tagger tasks encouraged the participants to consider the DDC suggestions if they thought them appropriate. While the figures should not therefore be considered a simple measure of popularity, in the questionnaires the participants commented favourably on suggestions in 26 different instances. They understood the benefit of having controlled tags (18 comments) in terms of ensuring consistency (e.g., 'allows you to tag along similar lines'), providing guidance and focus (e.g., 'useful to have different suggestions guiding in the right direction'), providing new suggestions (e.g., 'allowed you to include tags initially not thought of'), providing greater specificity (e.g., 'it allowed you to click on specific topics in much more detail'). Moreover, they appreciated the suggestions when they did not know the topic well (e.g., 'I like that if I was unsure about a subject in an article [Enhanced Tagger] usually hinted at good terms to use') and when they did not know which tags to assign (e.g., what one participant disliked about the Simple Tagger was '[l]ack of other suggestions if you could not decide upon appropriate tags'). They also appreciated automated suggestions for efficiency (e.g., 'saves time from typing').

4.2.1 Appropriateness of DDC

The DDC used in the demonstrator was 'raw DDC', i.e., DDC taken as is where its traditional purpose has been to group topically similar resources together. This instance of DDC is not necessarily appropriate to present online to end users for tagging. For example, a number of DDC captions contain quite complex phrases which are built terms. Others are very short or polysemic terms that are unclear without displaying their coordinate, superordinate or subordinate classes (in the demonstrator, displaying superordinate classes has been implemented through mousing over the class in the DDC treeview). Suggestions to deal with problematic terms for a future study or demonstrator, discussed in cooperation with OCLC, include the following (Vizine-Goetz and Panzer, 2008, personal communication):

1. Compressing long captions:
   - Abbreviate common terms (e.g., historical, geographical) and replace 'and' with ampersand;
   - Substitute the first relative index term for the caption (e.g., 'International party organizations' for 'International party organizations, auxiliaries, activities');
   - Truncate when the caption contains a series of terms (e.g., truncate 'International party organizations, auxiliaries, activities' into 'International party organizations…').
2. Captions that are unclear on their own: pre-coordinate them with a term from the DDC's Relative Index to provide context (e.g., the caption "Other conditions of employment" could be transformed into "Pensions and other conditions of employment").

A similar issue has been reported regarding the Bliss Bibliographic Classification 2 (Broughton, 2011) and its potential use a source for thesauri terms; many of the class names are not suitable as thesaurus terms and the formatting of class captions is not consistent, due to the lack of editorial policy in early schedules.

4.3 Retrieval implications

4.3.1 Tags versus metadata

In order to deduce the potential of the participants' tags to serve as additional access point in retrieval, the tags were compared against three fields from expert-created Intute metadata records: resource title, URL and description. These are the fields the participants could see when tagging so they could have been used as the source of tags. In addition, the tags were compared against controlled and uncontrolled keywords from the metadata records (both of which were hidden to the participants) because the major purpose of manually assigned keywords is to improve precision and recall in retrieval. While the term comparison was case-insensitive, spelling variations were not considered, i.e., each spelling variation was counted as a different term.

The analysis showed that end-user tags, both from the Simple and Enhanced Tagger, are found in 64% of tagged resources' title, URL or description. Thus, in comparison to those three parts of the metadata record which were visible to the participant, for 36% of the resources, tags potentially present additional access points. As seen from Table 6, most tags are found in the description (it being the longest piece of the metadata record), followed by the title, and then by the URL. Looking at the three fields, there are on average 5% less tags from the Enhanced Tagger found in the metadata records than there are from the Simple Tagger. In addition, out of 143 resources to which tags from DDC suggestions were assigned, the complete tags were found in only 44 resources (32 %), out of which 25 in description, 17 in title and 2 in URL. This suggests that, in retrieval, enhanced tagging could potentially provide more access points (that are not present in the resource text) than simple tagging when used in combination with a full text search engine. DDC terms which were found in titles and description were the short ones, e.g. DDC term 'peace' is found in title 'Multilateral peace operations database', 'Democratic Party' in 'Democratic Party Online (USA)', suggesting that longer DDC terms may potentially provide more access points, both as whole terms and for searching for individual words that are part of those terms.

Table 6. Number of resources' records where tags found

	Simple Tagger		**Enhanced Tagger**		**Total**	
Tags in title	200	48.3%	174	41.8%	350	44.9%
Tags in URL	76	18.4%	72	17.3%	137	17.6%
Tags in description	262	63.3%	232	55.8%	468	60.0%
Tagged resources in total	414		416		780	

Table 7 shows the degree to which simple and enhanced tags overlap with pre-assigned Intute keywords. Both simple and enhanced tags show very low overlap. The top two rows show the degree to which tags are the same as controlled and uncontrolled keywords, 4.5% in total. Third and fourth row include the exact matches and also situations where tags are parts of keywords, in which case there is 7.5% overlap in total. This small overlap indicates that, in comparison to Intute keywords, more than 90% of tags could serve as additional access points beyond the original Intute indexing.

Table 7. Tags versus pre-assigned (un)controlled keywords

	Simple Tagger		**Enhanced Tagger**		**Total**	
Tags same as controlled keywords	125	3.1%	119	3.4%	244	3.2%
Tags same as uncontrolled keywords	60	1.5%	39	1.1%	99	1.3%
Tags parts of controlled keywords	251	6.2%	175	4.9%	426	5.6%
Tags parts of uncontrolled keywords	84	2.1%	60	1.7%	144	1.9%
In total	520	12.9%	393	11.1%	913	12%

Of the tags selected from DDC, only a few overlap with existing Intute controlled keywords (these came from IBSS and HASSET thesauri): out of 598 DDC tags, 12 are the same as controlled keywords, and 5 the same as uncontrolled keywords. The following DDC tags are exactly the same as:

1. Controlled keywords: foreign policy, civil-military relations, intelligence, local government, democracy, diplomacy, nationalism, Rwanda, socialism, maps;
2. Uncontrolled keywords: Afghanistan, Liberia, Iraq, Iran, Hungary.

These are again relatively short DDC terms, also suggesting that longer DDC terms may potentially provide more access points, both as whole terms and for searching for individual words that are part of those terms. Also, end-user point of view may be different than that of Intute indexers, and providing another view is yet another benefit of social tagging.

Table 8 lists the results of qualitative comparison of simple and enhanced tags against titles and (un)controlled keywords, derived from the 12 selected resources (see section 3.2.5 for further explanation of the methodology and the match types. The results show that there are no significant differences across the two systems between title terms. There are slightly more preferred-term tags in the Enhanced Tagger, indicating that DDC suggestions may encourage taggers to control their tags, as the advantages of KOS have been reported by the participants. The synonym terms are only a few, and are mostly acronyms of their full form versions. There are twice as many acronyms in the Enhanced Tagger, implying perhaps that DDC suggestions inspire taggers to assign synonym variations. There are 4.1% more narrower tags in the Enhanced Tagger, indicating that the specificity of tags is higher when DDC suggestions are used. Inversely, the number of broader terms is slightly higher in the Simple Tagger in which the tags are generally broad (e.g., ‘security’, ‘law’). The largest number of tags in both systems are related-term tags. Whereas most related tags are topical, there are also tags denoting the resource type (e.g., ‘EU e-journals’), source/creator (e.g., ‘Kermit Blank’), audience/opinion (e.g., ‘excellent website for research’). There is about the same percentage of spelling mistakes in both systems. The type of spelling mistakes is, however, different in that in the Enhanced Tagger there are more spelling mistakes in names of organizations (‘Minstry of Defense’) or geographic names (e.g., ‘Afganistan’). This is surprising because geographic name tags are present in DDC and could have been selected. One recommendation is to use a KOS for automated spelling correction suggestions.

Table 8. Qualitative analysis of tags

Match types	**Simple system**	**Enhanced system**
Title term	71 (10.0%)	61 (9.9%)
Preferred term	21 (3.0%)	27 (4.4%)
Synonym term	8 (1.1%)	17 (2.7%)
Narrower term	126 (17.7%)	135 (21.8%)
Broader term	36 (5.1%)	28 (4.5%)
Related term	438 (61.6%)	339 (54.9%)
Spelling mistakes	11 (1.5%)	11 (1.8%)

Total	711 (100%)	618 (100%)

4.3.2 Search terms

One possible measure of indexing quality can be defined as the level of indexer-searcher consistency, i.e., the degree to which the tags assigned by taggers to a given resource match the tags used as search terms by searchers interested in that resource (Furner, 2010, p. 1864-1865). We could not do exactly that as this would require an additional set-up for which we lacked resources; what we were able to do is compare tags and pre-assigned keywords against the search terms the participants used in their tasks.

Table 9 compares tags and pre-assigned keywords to search terms which the participants used in their tasks. There were 98 search terms used in the Simple Tagger, and 122 in the Enhanced Tagger. In the Simple Tagger, almost three times as many search terms are found in tags as in controlled keywords, and twice as many as in uncontrolled keywords. The difference is even bigger for the Enhanced Tagger. The hypothesis that there is a greater probability of finding search terms in tags is an interesting avenue of future work. However, caution must be exercised since the study context might naturally encourage the use of search terms in tags, since tagging was always preceded by a search.

Table 9. Search terms versus tags and keywords

	Simple Tagger Number of search terms=98	**Enhanced Tagger** Number of search terms =122
Number of tags in which search terms are found	249	254
Number of controlled keywords in which search terms are found	91	40
Number of uncontrolled keywords in which search terms are found	111	101

There are 100 unique search terms contained in assigned tags. Out of these, 70 search terms are the same as tags (inclusive of 'peacekeeping' and 'European integration' from the controlled tasks) out of which the majority (40) are two-word terms (e.g., 'Barack Obama', 'Blair years', 'China trade', 'Cold War', 'Cuban history', 'debt forgiveness'). One-word terms are fewer (23) (e.g., 'democratisation', 'development', 'elections', 'Germany', 'globalization') and there are some three-word terms (7) (e.g., 'American foreign policy', 'conflict sensitive development', 'Cuban missile crisis', 'local strategic partnerships').

Of the 100 unique search terms, 6 are the same as assigned DDC tags: 'Afghanistan', 'democracy', 'Democratic Party', 'humanitarian intervention', 'international relations', 'Labour Party'. There are 64 search terms which are contained in assigned DDC tags. For example, a search using 'Afghanistan' as the search term retrieved a document which was then tagged a DDC term 'Afghanistan – Politics and government – 1973'; a search term 'Cuban missile crisis' assigned to a retrieved document DDC term 'Cuban Missile Crisis, 1962 – Sources'.

4.4 Post-task and post-study questionnaires

4.4.1 Post-task questionnaires

After each task, the participants completed a post-task questionnaire. It contained six questions with a five-point scale answers on topic familiarity, easiness to decide on tag choice, satisfaction with tags the participant assigned, certainty that the tags she assigned are correct, and usefulness of various features of the two interfaces in assisting them to choose tags. There was also space for comments on any other aspects that they thought important for the study (open-ended question).

As seen from the Table 10 below in both the Simple and Enhanced Tagger the majority of the participants were on average familiar with the topic of the task, they found it easy to choose tags, they were satisfied with tags they assigned and certain that they assigned the tags correctly. They found the following features useful: list of own tags, DDC treeview and DDC suggested tags. Main tag cloud was considered neither useful nor not useful; and so were others' names and DDC hierarchy. Others' names were considered not useful in free tasks which can in part be due to the fact that there were only 28 taggers and each chose a different free task.

Table 10. Post-task questionnaire with averaged values (5 representing highest, 1 lowest point)

Question	Simple Tagger		Enhanced Tagger		
	Controlled task	**Free task**	**Controlled task**	**Free task**	**Total**
Topic familiarity	3 (neutral)	4 (familiar)	4 (familiar)	4 (familiar)	4 (familiar)
Easiness to choose tag	4 (easy)	4 (easy)	4 (easy)	4 (easy)	4 (easy)
Satisfaction with assigned tags	4 (satisfied)	4 (satisfied)	4 (satisfied)	4 (satisfied)	4 (satisfied)
Certainty in correctness of assigned tags	4 (certain)	4 (certain)	4 (certain)	4 (certain)	4 (certain)
Usefulness of main tag cloud	3 (neutral)	3 (neutral)	3 (neutral)	3 (neutral)	3 (neutral)
Usefulness of list of own tags	4 (helpful)	4 (helpful)	4 (helpful)	4 (helpful)	4 (helpful)
Usefulness of others' names	3 (neutral)	2 (unhelpful)	3 (neutral)	2 (unhelpful)	3 (neutral)
Usefulness of DDC treeview			4 (helpful)	4 (helpful)	4 (helpful)
Usefulness of DDC hierarchy			3 (neutral)	3 (neutral)	3 (neutral)
Usefulness of DDC suggested tags			4 (helpful)	4 (helpful)	4 (helpful)

Detailed analysis additionally shows that when the participants knew the topic well and tagged in the Simple Tagger, they found it very easy (5) to choose a tag and were more satisfied and more certain about the tag accuracy than in other task-interface combinations. Thus, topic familiarity seems to help tagging.

As part of the post-task questionnaire, there was also space for comments on any other aspects that they thought important for the study. The comments on the Simple Tagger related to the following (number of times identified in parentheses):

- tagging (6): they identified the problem of synonyms (2), found it easier to tag when the topic was familiar (1), assigned fewer tags when topic was familiar as one knows what is important (1), they were not entirely sure of tags assigned (1), and, they found it quicker to tag in the Simple Tagger (1)
- interface (5): they recognized a software error (2), found it quicker to tag in the Simple Tagger because the Enhanced Tagger interface is not streamlined (1), they considered main tag cloud visually unclear it taking a lot of time to scroll through (1), they thought it complex to get around the interface (1)
- Intute related (4): they recognized the search-engine issue of retrieving irrelevant search results (2), they minded Intute's orientation to UK resources (1), they found useful resources (1).

The comments on the Enhanced Tagger related to the following (number of times identified in parentheses):

- interface (7): there is a need for better presentation of DDC suggestions (2), other existing tags would be useful (1), there is a poor viewing ability in the demonstrator (1), main tag cloud is very large and it is difficult to scroll through (1), the upload is slow (1), and, when reloading the pointer goes back to the top of the page and you need to scroll down again (1)
- tagging (7): DDC terms are variable in accuracy – sometimes they are completely irrelevant, sometimes spot-on (4), there are no DDC suggestions for some topics (e.g., human trafficking, Tony Blair) (2), and, DDC is United States centric (1)
- Intute (4): there are search engine issue (some search results are missing) (2), a suggestion to include Spanish (1), and, there are dead links (1)
- general (1): 'Enhanced Tagger not that great' (1).

These comments point to interface issues such as its complexity and the need to scroll around a lot, and more so in the Enhanced Tagger, including the Main tag cloud presentation. Related to tagging, some thought the Simple Tagger quicker, but wondered how to deal with synonyms such as acronyms and different word forms. In the Enhanced Tagger, DDC suggestions proved variable in accuracy, with no suggestions whatsoever for certain topics.

4.4.2 Post-study questionnaire

At the very end of the study, the participants filled in a post-study questionnaire about their experience as a whole. It comprised five five-point scale questions relating to the degree to which they enjoyed the study

and the degree to which they thought it easy to learn and use the two interfaces. Another question examined the usefulness of an analogues system in real life. Four open-ended questions related to what they liked and what they disliked about the two interfaces. The final question asked for general comments and suggestions.

Table 11. Post-study questionnaire with averaged values

Question	Scale	Answers
Enjoying the study	5 (totally)… 1 (not at all)	3 (kind of)
Easy to learn Simple Tagger	5 (extremely)… 1 (not at all)	4 (very)
Easy to use Simple Tagger	5 (extremely)… 1 (not at all)	4 (very)
Easy to learn Enhanced Tagger	5 (extremely)… 1 (not at all)	3 (somewhat)
Easy to use Enhanced Tagger	5 (extremely)… 1 (not at all)	3 (somewhat)
Useful in real life	3 (yes) 2: (I don't know) 1 (no)	3 (yes)

As seen from Table 11, the majority enjoyed the study to a medium degree. They thought it very easy to learn and very easy to use the Simple Tagger, and somewhat easy to learn and somewhat easy use the Enhanced Tagger. The majority thought a similar system would be useful in real life (14) as an an organized system for researching relevant resources, for team projects, for specific research topics, and when starting out a new project. They said that it would make search results more specific and that tagging makes it easier to find research resources and that it is more helpful than browsers with favourites. Some praised good quality resources in the collection. Some thought it would be useful but under certain conditions (8) including better suggestions, a simplified Enhanced Tagger and easier navigation, advanced search suggestions, having dealt with the challenge of how to make people to tag as well as including more academic discourse and having more resources included Four thought it not useful in real life stating that for retrieval search engines are more effective and that there is no point in tagging as title says it all.

The comments about the Simple Tagger show that the participants liked the Simple Tagger for having their own choice (10), ease of use (7), it being quicker to tag than in the Enhanced Tagger (5), simplicity of tagging (5), simplicity of the interface (2), it being better when one knows the topic and which tags to use (2), and it being less distracting than suggestions from the Enhanced Tagger and it having less influence from other tags (2). They disliked it mostly for reasons related to tagging (11) such as lacking suggestions when e.g., not knowing which tags to assign in a non-familiar topic and having the problem of synonyms, poor layout (5) such as need for scrolling, large main tag cloud etc. and technical and search engine issues (5) mentioned above. They also considered interface could be improved including usability of the Main tag cloud.

The comments about the Enhanced Tagger show that the participants appreciated the KOS suggestions for reasons such as terms not initially thought of, ensuring consistency in the collection, hinting at terms not initially thought of and finding suitable terms when unfamiliar with the topic as well as providing more specific terms (16 comments). However, the suggestions were sometimes inadequate (16), and even absent for some topics (4). Also, they did not like the complex interface (6) and listed some technical issues (5). While some thought the Enhanced Tagger to be time-saving (2), others commented it took longer to input tags (3).

The general comments and suggestions more or less repeat the participants' previous input. Most of the suggestions relate to improving the interface (8) and the DDC tags (2).

5 Conclusions

In order to compare simple tagging as common in existing applications against tagging enhanced with suggestions from a KOS (DDC with LCSH mappings), an exploratory user study was conducted. It investigated indexing aspects and influence on retrieval when using only social tagging versus when using social tagging in combination with the KOS. The study involved a demonstrator with two interfaces, the Simple Tagger and the Enhanced Tagger, both having features common in tagging applications and the latter with the additional feature of automated suggestions from the KOS. There were 28 participants who each tagged 60 resources across 4 different tasks, 2 of which were on a controlled topic with a group task

scenario, and 2 were on a freely chosen topic; 2 were completed in the Simple Tagger and 2 in the Enhanced Tagger. The data collection included over 11,000 Web-based textual resources in the area of politics. The analysis was based on the log data of actions the participants took including the log analysis of user sessions, and on a set of three types of questionnaires.

In both the Simple and Enhanced Tagger, the majority of the participants were on average familiar with the topic of the task, they found it easy to choose tags, were satisfied with tags they assigned, and were certain that they assigned the tags correctly. They assigned over 7500 tags in total. More tags were assigned in the Simple Tagger and in free tasks. The former could be explained by the fact that exploring KOS suggestions in Enhanced Tagger requires more effort and hence as a result fewer tags are assigned, and the latter by more intrinsic motivation. Each resource received on average 5 tags, which is similar to findings about real-life tagging applications. The distribution of tags follows the power law in both interfaces, which is also commonly reported in other studies.

In the Simple Tagger the majority of tags are assigned by typing them directly in, again as common in social tagging applications In the Enhanced Tagger 20% less tags are typed and 17% of those come from DDC-based suggestions. Of other sources, in both interfaces tags mostly come from another tagger's tag and the main tag cloud. The questionnaires show that the participants found the list of own tags, DDC treeview and DDC suggested tags useful, whereas main tag cloud, others' names and DDC hierarchy were considered neither useful nor not useful. Others' names were considered not useful in free tasks which is understandable because there were only 28 taggers and each chose a different free task. The main tag cloud became cumbersome in time and required a lot of scrolling.

The instructions for the Enhanced Tagger tasks encouraged the participants to consider the DDC suggestions if they thought them appropriate; however, the questionnaires imply the participants appreciated the suggestions in particular because they provided tags the participant did not initially think of and as the controlled tags would ensure consistency in the collection. However, the suggestions were sometimes considered inadequate, and even absent for some topics. The DDC used in the demonstrator was DDC as is, and as such it is not necessarily appropriate to present online to end users for tagging. For example, a number of DDC captions contain quite long phrases, others require hierarchical context etc. and need to be adjusted for the purpose of the end-user indexing and searching.

Implications for retrieval are that in comparison to existing resource description, title and URL fields of metadata records, tags from both interfaces present potential additional access points for 36% of the resources, and a bit more so for the DDC tags. In comparison to Intute controlled and uncontrolled keywords, more than 90% of tags could serve as additional access points beyond the original Intute indexing. Moreover, in the Simple Tagger, almost three times as many search terms are found in tags as in controlled keywords, and twice as many as in uncontrolled keywords. The difference is even bigger for the Enhanced Tagger. Qualitative comparison of tags in the Simple and in the Enhanced Tagger against title and (un)controlled keywords shows that there are more of what we defined as preferred term tags, synonymous tags and narrower term tags in enhanced than in the simple tags, all of which implies the potential for increased precision and recall based on the Enhanced Tagger, which is useful in web searching in large collections.

The majority of the participants thought a similar system would be useful in real life for research purposes as an organized system for resources; some thought it would be useful in real life if suggestions were better and if interface were improved. A few considered that search engines were more effective. Based on the participants' input, the improvements should focus on the: 1) interface, making it less complex for use in general and less complex for DDC exploration in particular with less scrolling, and 2) DDC suggestions including improved automated suggestion algorithm and adjustment of DDC terms for end-user presentation and use in social tagging.

Apart from interface and DDC adjustments, further research could also include automated spelling correction and tag form unification. Different KOS and how to adjust them for social tagging should also be investigated. The motivation of the participants behind tagging and their cognitive processes should be investigated. All this would be conducted in the context of a comprehensive retrieval test.

Overall, the study indicates that tagging enhanced with suggestions from DDC or another well-established knowledge organization system can make important improvements to the usual simple tagging, especially if the improvements suggested in this paper are implemented. This refers in particular to enhanced tags providing more precise or specific index terms which is what searchers commonly use in web searching

today. Even if taggers are not professional indexers, the value gained by providing suggestions from an established KOS such as DDC is considerable, assuming that the taggers' motivation is appropriate and altruistic. The issue of what might be the best compromise between the costs of imposing control through professional KOS-based subject indexing versus the ambiguity and other problems of free tags is a complex question depending on numerous factors and contexts. In our view, enhanced tagging does make possible the indexing of subject aspects that may not be covered by professional indexers or automated subject indexing tools.

6 References

Angus, E., Thelwall, M., and Stuart, D. (2008), "General patterns of tag usage among university groups in Flickr", *Online Information Review*, Vol. 32 No. 1, pp. 89-101.

Ann Arbour District Library. (2011), "Catalog", available at: http://www.aadl.org/catalog (accessed 23 May 2011)

Broughton, V. (2011), "Facet analysis as a tool for modelling subject domains and terminologies", in Slavic, A., Civallero, E. (Eds.), *Classification and ontology: formal approaches and access to knowledge: proceedings of the International UDC Seminar*, 19-20 September 2011, The Hague, The Netherlands, Würzburg, Ergon Verlag, pp. 207-228.

Catarino, M.E. and Baptista, A.A. (2008), "Relating folksonomies with Dublin Core", paper presented at the International Conference on Dublin Core and Metadata Applications, 22-26 September 2008, Berlin, available at: http://dcpapers.dublincore.org/ojs/pubs/article/view/915/911 (accessed 23 May 2011)

Del.icio.us. (2011), available at http://www.delicious.com/ (accessed 6 June 2011)

ExLibrs. (2011), "Primo: A new user experience", available at: http://www.exlibrisgroup.com/category/PrimoUserExperience (accessed 23 May 2011)

Flickr. (2011) http://www.flickr.com/ (accessed 6 June 2011)

Golub, K., Jones, C., Lykke Nielsen, M., Matthews, B., Moon, J., Puzoń, B., and Tudhope, D. (2009a), "EnTag: Enhancing Social Tagging for Discovery", in Heath, F., Rice-Lively, M. L., Furuta, R. (Eds.), *Joint Conference on Digital Libraries (JCDL), Austin, TX, June 15-19,* ACM, New York, NY, pp. 163-172.

Golub, K., Jones, C., Lykke Nielsen, M., Matthews, B., Moon, J., and Tudhope, D. (2008a), "Enhancing social tagging with a knowledge organization system", *Aliss Quarterly*, Vol. 3 No. 4, pp. 13-16.

Golub, K., Lykke Nielsen, M., Moon, J., and Tudhope, D. (2009b), "Enhancing social tagging with a knowledge organization system", paper presented at IFLA 2009 Satellite Meeting "Emerging trends in technology: libraries between Web 2.0, semantic web and search technology", 19-20 August 2009, Florence, available at: http://homes.ukoln.ac.uk/~kg249/publ/entag-ifla-v3-final.pdf (accessed 1 May 2013)

Furner, J. (2010), "Folksonomies", Bates, M., *Encyclopedia of library and information sciences*, 3rd ed., CRC Press, New York, pp. 1858-1866.

Golub, K., Lykke Nielsen, M., and Tudhope, D. (2008b), "Enhanced Tagging for Discovery (EnTag): User study", available at: http://www.ukoln.ac.uk/projects/enhanced-tagging/study/ (accessed 1 May 2013)

Hayman, S. (2007), "Folksonomies and tagging: New developments in social bookmarking", paper presented at the Ark Group Conference: Developing and Improving Classification Schemes, 27-29 June, Sydney, available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.138.8884 (accessed 1 May 2013)

Innovative Interfaces. (2011), “Encore”, available at: http://www.iii.com/products/encore.shtml (accessed 23 May 2011)

Intute Consortium. (2010), “Intute”, available at: http://www.intute.ac.uk/ (accessed 7 June 2011)

Lykke, M., Høj, A. L., Line Nørgaard, M., Golub, K., Tudhope, D. (2011), “Tagging behaviour with support from controlled vocabulary”, in Gilchrist, A. and Vernau, J. (Eds.), *Facets of knowledge organization: A tribute to Professor Brian Vickery 1918-2009, Proceedings of the ISKO UK second biennial conference*, London, 4th-5th July, pp. 41-50.

Marchetti, A., Tesconi, M., Ronzano, F., Rosella, M., and, Minutoli, S. (2007), “SemKey: A semantic collaborative tagging system”, paper presented at Workshop on Tagging and Metadata for Social Information Organization, 16th International World Wide web Conference, May 8-12, Banff, Aberta, Canada, available at: http://www2007.org/workshops/paper_45.pdf (accessed 1 May 2013)

Montana State University. (2011), “Electronic thesis and dissertations (ETDs)”, available at: http://etd.lib.montana.edu/etd/view/ (accessed 23 May 2011)

Moon, J., Tudhope, D., Golub, K., Lykke Nielsen, M. (2008), “Enhanced Tagging for Discovery (EnTag): Demonstrators”, available at: http://hypermedia.research.glam.ac.uk/kos/entag/ (accessed 1 May 2013)

Naaman, M., and Nair, R. (2008), “ZoneTag's collaborative tag suggestions: What is this person doing in my phone?”, *IEEE Multimedia*, Vol. 15 No. 3, pp. 34-40, available at: http://infolab.stanford.edu/~mor/research/ZoneTagIEEEMM.pdf (accessed 1 May 2013)

National Information Standards Organization (2005), *Guidelines for the construction, format, and management of monolingual controlled vocabulares, ANSI/NISO Z39.19-2005*, NISO Press, Bethesda, MD.

National Library of Australia. (2010), “Picture Australia: Flickr”, available at: http://www.pictureaustralia.org/contribute/participants/Flickr.html (accessed 23 May 2011)

Noruzi, A. (2007), “Folksonomies: Why do we need controlled vocabulary? Editorial”, *Webology*, Vol. 4 No. 2, available at http://www.webology.ir/2007/v4n2/editorial12.html (accessed 1 May 2013)

OCLC. (2011), “DDC services”, available at: http://www.oclc.org/DDC/ (accessed 7 June 2011)

Panteion University. (2008), “Opacial”, available at: http://library.panteion.gr/opacial/ (accessed 23 May 2011)

ProQuest. (2011), “IBSS: International Bibliography of the Social Sciences”, available at: http://www.proquest.co.uk/en-UK/catalogs/databases/detail/ibss-set-c.shtml (accessed 7 June 2011)

Raymond, M. (2008), “My friend Flickr: A match made in photo heaven”, Library of Congress Blog, available at: http://www.loc.gov/blog/?p=233 (accessed 30 May 2011)

Rolla, P. J. (2009), “User tags versus subject headings: Can user-supplied data improve subject access to library collections?”, *Library Resources & Technical Services*, Vol. 53 No. 3, pp. 174-184.

Spiteri, L. F. (2007), “Structure and form of folksonomy tags: The road to the public library catalogue”, *Webology*, Vol. 4 No. 2, available at: http://www.webology.org/2007/v4n2/a41.html (accessed 23 May 2011)

Syn, S. Y. and Spring, M. B. (2013), “Finding subject terms for classificatory metadata from user-generated social tags”, Journal of the American Society for Information Science and Technology, Vol. 64, pp. 964-980.

UK Data Archive. (2011), “Our HASSET thesaurus”, available at: http://www.data-archive.ac.uk/find/hasset-thesaurus (accessed 7 June 2011)

University of Pennsylvana. (2005), “PennTags: Help”, available at: http://tags.library.upenn.edu/help/ (accessed 23 May 2011)

Weller, K., (2007), “Folksonomies and ontologies: Two new players in indexing and knowledge organization”, in Jezzard, H. (Ed.), *Applying Web 2.0: Innovation, Impact and Implementation, Proceedings of the Online Information Conference*, Learned Information Europe, London, pp. 108-115.

Yi, K., Chan, L. M. (2009), “Linking folksonomy to Library of Congress subject headings: An exploratory study”, *Journal of Documentation*, Vol. 65 No. 6, pp. 872-900.